PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: **H04M 15/00, G06F 15/18, H04L 12/46**	A2	(11) International Publication Number: **WO 00/45584** (43) International Publication Date: 03 August 2000 (03.08.2000)

(21) International Application Number: PCT/US00/02011

(22) International Filing Date: 28 January 2000 (28.01.2000)

(30) Priority Data:
60/117,663 28 January 1999 (28.01.1999) US

(60) Parent Application or Grant
BIOS GROUP LP [/]; (). KAUFFMAN, Stuart, A. [/]; (). PLATE, Tony, A. [/]; (). KAUFFMAN, Stuart, A. [/]; (). PLATE, Tony, A. [/]; (). MORRIS, Francis, E. ; ().

Published

(54) Title: **A METHOD AND SYSTEM FOR ROUTING CONTROL IN COMMUNICATION NETWORKS AND FOR SYSTEM CONTROL**

(54) Titre: **PROCEDE ET SYSTEME D'ACHEMINEMENT DE COMMANDES DANS DES RESEAUX DE COMMUNICATION, ET DE COMMANDE DU SYSTEME**

(57) Abstract

The present invention relates generally to a method and system for routing control in communication networks and for system control. More particularly, the present invention performs routing by controlling the components in a network with software agents (102) operating in a reward framework using p, tau, and patches (104) to improve communication performance (106). This invention disclosure includes the combination of reinforcement learning agents in a market-based or performance-based reward framework together with optimization techniques called p, tau, and patches (104) as applied to the problem of topology-and load-based routing in data networks, in order to improve communication performance (106) such as communication latency and bandwidth. The invention also applies to the control of other systems, including operations management, job-shop problems, organizational structure, portfolio management, risk management etc.

(57) Abrégé

La présente invention concerne, de manière générale, un procédé et un système d'acheminement de commandes dans des réseaux de communication, et de commande du système. Pour réaliser les acheminements, la présente invention gère, en particulier, les composants du réseau à l'aide d'agents (102) logiciel travaillant en réseau à récompense utilisant les fonctions p, tau, et patches (104) de façon à améliorer le rendement (106) des échanges. Le procédé selon l'invention prend des agents d'apprentissage d'un réseau à récompense orienté marché ou rendement et les associe à des techniques d'optimisation de type p, tau, et patches (104), que l'on utilise pour résoudre les problèmes posés par les acheminements de type topologie et charge dans le cas des réseaux de données, et ce de façon à augmenter le rendement (106) des échanges, notamment les points tels que les mises en attente et les largeurs de bande. L'invention concerne également la gestion des autres systèmes, notamment la gestion de la mise en oeuvre, les problèmes de gestion des travaux, la structure d'organisation, la gestion des actifs, la gestion des risques, etc.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : H04M 15/00, H04L 12/46, G06F 15/18

A2

(11) International Publication Number: WO 00/45584

(43) International Publication Date: 3 August 2000 (03.08.00)

(21) International Application Number: PCT/US00/02011

(22) International Filing Date: 28 January 2000 (28.01.00)

(30) Priority Data:
60/117,663 28 January 1999 (28.01.99) US

(71) Applicant *(for all designated States except US)*: BIOS GROUP LP [US/US]; 317 Paseo de Peralta, Santa Fe, NM 87501 (US).

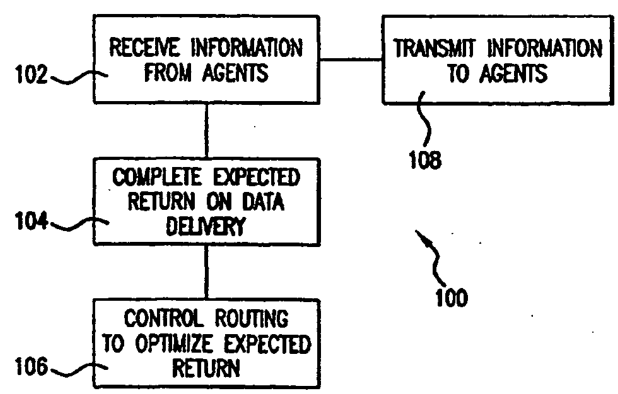
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: KAUFFMAN, Stuart, A. [US/US]; 1811 S. Camino Cruz Blanco, Santa Fe, NM 87505 (US). PLATE, Tony, A. [US/US]; 812 Loma Boreal, Santa Fe, NM 87591 (US).

(74) Agents: MORRIS, Francis, E. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: A METHOD AND SYSTEM FOR ROUTING CONTROL IN COMMUNICATION NETWORKS AND FOR SYSTEM CONTROL



(57) Abstract

The present invention relates generally to a method and system for routing control in communication networks and for system control. More particularly, the present invention performs routing by controlling the components in a network with software agents (102) operating in a reward framework using *p*, *tau*, and *patches* (104) to improve communication performance (106). This invention disclosure includes the combination of reinforcement learning agents in a market-based or performance-based reward framework together with optimization techniques called *p*, *tau*, and *patches* (104) as applied to the problem of topology-and load-based routing in data networks, in order to improve communication performance (106) such as communication latency and bandwidth. The invention also applies to the control of other systems, including operations management, job-shop problems, organizational structure, portfolio management, risk management etc.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

A METHOD AND SYSTEM FOR ROUTING CONTROL IN COMMUNICATION NETWORKS AND FOR SYSTEM CONTROL

FIELD OF THE INVENTION

The present invention relates generally to a method and system for routing control in communication networks and for system control. More particularly, the present invention performs routing by controlling the components in a network with software agents operating in a reward framework using *p*, *tau*, and *patches* to improve communication performance.

Background

Modern data-communication networks, as a non-limiting example packet-switched data networks, often present many potential routes between nodes that wish to communicate. Decisions about the route that data should take are usually made in a decentralized fashion by routers at the nodes. Decisions must be decentralized both because a centralized routing device would make the network vulnerable to single-point failures and because it would be impractical to communicate routing decisions from a centralized device to all the nodes in a spatially disperse network. Ideally, routing decisions should take into account both network topology (e.g., finding the shortest or least-cost path between two nodes) and current and historical network load (i.e., finding paths that do not utilize currently or historically overloaded communication links).

However, it is difficult to construct routers that make effective decisions based on load due to the problem of oscillation. For example, if link A is currently overloaded and link B is currently under loaded, then link B appears preferable to all the routers, which leads to link B being overloaded and link A being under loaded, and so on. Consequently, currently-fielded commercially-available routers take into account only network topology when making

routing decisions (though they may try to split traffic among equal-cost paths.) As a result, communication performance is not as good as is theoretically possible. Bandwidth, delay (latency) and reliability (i.e., packet loss) are all negatively affected by routing decisions that do not take network load into account.

Accordingly, there is a pressing need for decentralized routing algorithms that can effectively take both network topology and current and historical load on communication links into account.

Summary of the Invention

The present invention present a method and system for routing control in communication networks by controlling the components in a network with software agents operating in a reward framework using *p*, *tau*, and *patches* to improve communication performance.

The present invention includes a method for routing packets of data through a network of a plurality of components comprising the steps of:

controlling one or more of said components by executing a corresponding one or more software agents, comprising the steps of:

receiving information for at least one of the packets;

computing an expected return for delivery of said at least one packet from said information; and

directing the delivery of said at least one packet to optimize said expected return.

The present invention includes a method for routing packets of data through a network of a plurality of components comprising the steps of:

defining at least one algorithm having one or more parameters for routing the data;

defining at least one global performance measure of said at least one algorithm;

executing said algorithm for a plurality of different values of said one or more parameters to generate a corresponding plurality of values for said global performance measure;

constructing a fitness landscape from said values of said parameters and said corresponding values of said global performance measure; and

optimizing over said fitness landscape to generate optimal values for said at least one parameter.

Brief Description of Drawings

FIG. 1 provides a flow diagram describing the operation of software agents that direct the delivery of packets of data by controlling corresponding components in a communication network.

FIG. 2 provides a flow diagram for determining optimal values of parameters of methods performing routing control and system control.

Detailed Description of the Preferred Embodiment

The present invention consists of installing an independent software agent at one or more routers. In the preferred embodiment, the independent software agents are installed in some or all of the routers at any level in a hierarchy of networks and subnetworks. Each software agent updates the routing information (as a non-limiting example, routing tables) in the memory of its associated router, and shares connectivity and load information with other software agents. The software agent may either run on the same processor as its associated router or on a different processor.

Each agent acts autonomously to optimize the value of some function combining its own performance index, and that of some (zero or more) selected neighbors (not necessarily immediate topological neighbors) as explained more fully below. The performance index is based on one of the following:

(a) its "earnings" from transmitting packets; or

(b) a local measure of communication performance such as combining indices of load on adjacent links and expected delivery times of packets passing through its router.

Agents learn to optimize their performance index using reinforcement learning. An exemplary reinforcement learning technique is Q-learning.

Without limitation, the following embodiments of the present invention are described in the illustrative context of a solution that installs software agents at the routers of a communication network. However, it will be apparent to persons of ordinary skill in the art that the present invention also applies to the use of software agents to

control other components of the communication network. For example, software agents could control one or more directional or non-directional communication links.

FIG. 1 provides a flow diagram 100 describing the operation of software agents that direct the delivery of packets of data by controlling corresponding components in a communication network. In step 102, the software agent receives information on a packet of data from other software agents. Next, in step 104, the software agent computes an expected return for delivering the packet of data using the information. Next, in step 106, the software agent controls the routing of the data through its corresponding component to optimize the expected return. In step 108, the software agent transmits information to other software agents so that they can similarly control their corresponding components to optimize their expected return.

Integration With existing technology

As a non-limiting example, the present invention integrates with existing standards surrounding the Open Shortest Path First (OSPF) routing standard (RFC-2328) as follows:

Routing tables for OSPF-compatible routers : Preferably, the agents will not make routing decisions for each and every communication request. For example, the software agents will not make routing decisions for each packet that is to be routed towards some destination. Instead, the agents will modify the routing information that the routing software or hardware uses to make decisions about communication requests. Preferably, the routing information is stored in routing tables. Thus, the agent may take a significant amount of time to perform a single action such as changing one entry in a routing table. Further, this single action may

subsequently affect decisions made by the router for an indefinite period of time.

Hash-based load division : As a non-limiting example, in packet-switched networks it is usually desirable to route all packets from the same source destined for the same destination along the same route. This scheme is used to prevent out-of-order arrival of packets. This scheme can be accomplished in OSPF-compatible routers by partitioning packets for the same destination host or subnet into classes based on a hash function of the source and destination host network addresses. The classes are contiguous regions of the range hash function and the borders of these regions are defined by the routing tables. The hash value could also be a function of other packet header parameters such as a reward value and quality of service specifications as defined in detail below.

Opaque Link State Advertisements : Agents must be able to communicate information about local topology and load to other agents. Preferably, this information is in the form of bids for the delivery of packets. Alternatively, this information may be directly encoded. The communication of this information takes priority over regular data traffic in the network in order to ensure its timely arrival at nodes where it is needed. As a non-limiting example, this information could be packaged in Opaque Link State Advertisements packets (RFC--2370).

Hierarchical network structure : Networks may be structured hierarchically such that the internal structure of subnets are only visible from within the network. It will be apparent to persons of ordinary skill in the art that the present invention applies to all schemes that can be used in hierarchical networks with the modification that some of the entries in the routing tables cover groups of destinations. Similarly, some of the bids are for groups of destinations.

Agent performance indices

Agents receive immediate feedback about their performance. This feedback is called a *reward*. However, in the reinforcement learning framework of the present invention, an agent does not merely act to optimize its immediate reward. Instead, it acts to optimize its *return*. In the preferred embodiment, the *return* includes an expected future reward that is discounted to present value. As mentioned earlier, reward is based on "earnings" in a communication market in one of the preferred embodiments called the market-based reward framework. In another preferred embodiment called the local performance reward framework, the reward is based on an index of local communication performance.

Market-based reward framework

In the market-based reward framework, each packet contains a contract to pay some amount of a "cash" equivalent to the router that delivers it to its final destination. The contracted amount is paid in full only if the packet reaches its final destination within a constraint such as a pre-

specified quality of service constraint. Preferably, a portion of the contracted amount is paid at the destination if the packet arrives outside of the specified quality of service. This portion is determined as a function of the received quality of service. Preferably, less cash is released for packets that arrive with excessively long latency (for interactive connections). Likewise, less cash is released for packets that arrive out-of-order or at widely varying intervals (for audio or video streams). At the final destination node of a packet, market-arbiter software calculates the cash reward earned by the delivering software agent and the amount owed by the originating application. These rewards and bills are accumulated over time and sent out at a low frequency so as to impose only a negligible communication load on the network.

When reinforcement learning is used to adjust the behavior of agents, instantaneous rewards are based on the actual cash profit of the agent and optionally, the cash profit of neighboring agents (not necessarily topological neighbors) over some short past time period. Optionally, in order to prevent agents from charging arbitrary prices in monopoly situations, excess profit can be removed (taxed) from those agents whose long-term discounted expected reward exceeds a predefined target.

Each agent communicates "bids" that specify how much it will pay for packets having a particular destination, a particular specified quality of service, and a specified maximum rate to other agents. Preferably, each agent communicates the "bids" to its topologically neighboring agents. Bids may also have an expiration time. Optionally, the bids are represented by a function. Non-limiting function examples include a margin, a rate, a minimum contract value, and a minimum delivery time. For example, an

agent at node B may specify that it will pay the value less 3 units for up to 800 packets per second destined for node F having a value of at least 15 units and a remaining allowable delay of 120ms. Bids stand until they expire or until the node where a bid is held receives a message canceling and/or replacing the bid. Optionally, other quality of service parameters corresponding to the quality of service requirements of packets are included in the bids. For example, a higher price may be paid for packets that arrive in sequence. Bids may also specify a route. When bids specify a route, agent may not sell a packet against a bid that would result in the packet returning to the same router. For example, if B submits a bid to A to deliver packets to E via the path CDAF, then A may not sell to B packets destined for E.

Packets that are received by a node (either from an application program at the node, or from another node) that do not conform to the parameters of an existing bid (e.g., insufficient contract value or too many in a given time period) do not require payment. Instead, these packets are owned by the agent at the node and may be sold.

Optionally, in addition to the agent software, nodes also execute market-arbiter software. The market-arbiter software keeps track of bids and updates and allocates payment for packets in accordance with the previously discussed market rules. Optionally, bids specify "preference surfaces" that give propensities to buy or sell as probabilistic functions of qualify of service, delay, and other features. Preference surfaces were defined in co-pending patent application number 09/345,441, titled, "An Adaptive and Reliable System and Method for Operations Management" and filed on July 1, 1999, the contents of which are herein incorporated by reference. Preferably, the

market-arbiter software matches preference surfaces of bidders and sellers to optimize a total "utility" for a group of packets and routers.

Preferably, agents make decisions based on sources of information. The decisions include:

the determination of bids and bid updates to submit to other software agents, and

the modification of the routing tables to direct packet flow so as to optimize the expected return on the routed packets.

The sources of information include:

bids received from other agents,

measured flows of packets through the associated router of the agent, and

the expected return at the router and at neighboring routers (that are not necessarily neighbors in the topographical sense).

The execution of the software agents using these market rules lead to the following network behavior:

- *Agents will pay more for packets nearer the destination.* The agent in the destination node receives the contract value in the packet when it delivers the packet to the destination application. Consequently, it will be willing to pay a high price (near the contract value) for such packets. The agent in next-to-last node will be willing to pay a slightly lower price, and so on. Packets far from their destination will be purchased for relatively little.

- *It will generally cost more to send packets further.* Since the agent at each node along a route takes its own margin (e.g., buys packets for 8 units, and sells them for 10 units), it will cost more to send packets further. Preferably, the margins charged by agents

reflect actual establishment and/or operating costs for particular communication links.

- *Different levels of service may be provided.* An agent may maintain different bids for different levels of service. Higher levels of service such as a faster delivery time will cost more. A packet that is sent out with sufficient contract value to cover a higher level of service but that does not arrive at its destination within the specified quality of service parameter will only be worth a reduced value to the router making the final delivery. In this situation, the originating application will be charged only the reduced value.

- *Application programs at nodes will know how much it costs to send a packet to a particular destination.* The bids lodged at a node specify how much it costs to send a packet to a particular destination. Once the packet is in transit, even if routing costs change, intermediate nodes are still motivated to forward packets as explained further in the next paragraph.

- *Packets are always worth sending.* Even if an agent is caught in a crunch, it is still worthwhile for the agent to sell the packets at a loss. For example, suppose an agent receives 500 packets at a price of 7 units, expecting to be able to sell them for 9 units. Suppose further that the bid drops to 3 units before the agent can sell them. Even in this situation, the agent will sell the packets at a loss because if it retains these packets, it receives no reward at all from them as their contract value is not realized until they reach their destination.

- *Agents will have to make predictions about future packet flow.* Since decisions cannot be made about individual packets but only about bids and routing table entries, earnings will depend on the flow of packets and may fluctuate. Preferably, agents make predictions about future packet flow in order to set routing table entries so as to maximize expected return. For example, an agent may set routing table entries to forward most of the received packets to a neighbor who pays well for them (but not too many, since it will not receive a reward for the ones sold above a predetermined rate as explained in the preceding monopoly discussion).

- *Agents will be motivated to keep bids up-to-date and high.* If an agent charges too large a margin (ie., its bids are too low), it will loose business to competitors, and consequently will receive a lower return. If an agent lets its bids get out-of-date and too high, it will receive a lower or negative return on packets that it forwards. Hence, agents will be motivated to keep bids high (i.e. margins low) and up-to-date.

- *Earnings at nodes can help guide decisions about short- and long-term resource allocation.* If margins at nodes are designed to accurately reflect costs of communication, then market theory indicates that prices charged by agents will accurately reflect benefits of allocating additional resources (barring monopoly situations). Thus, prices charged by agents can be used as a guide for allocating short-term or long-term resources such as a temporary connection or a leased line.

Local-performance reward framework

An alternative to the market-based reward scheme is a scheme where local rewards are based on unbiased estimates of packet delivery times. Preferably, packet delivery times are estimated in a decentralized fashion by plugging reported link loads into models of network performance. The immediate reward for an agent at a node is the inverse of an increasing function of the aggregate estimated packet delivery times. Optionally, the immediate reward also incorporates other indices of quality of service. In the local performance reward framework, agents modify routing tables in an attempt to reduce the estimated delivery times or improve other aspects of quality of service.

Locally-cooperative local reinforcement learning

Having all agents attempt to optimize their local figures of merit will not always result in the discovery of the globally optimum configuration as explained in "At Home in the Universe" by Stuart Kauffman, Oxford University Press, Chapter 11 in the context of an NK fitness landscape, the contents of which are herein incorporated by reference. This result occurs because actions taken by one agent affects its state and possibly changes the context of the reward for its neighboring agents.

Accordingly, in the preferred embodiment the present invention utilizes combinations of the following three semi-local strategies:

patches In this technique, agents are partitioned into disjoint subsets called patches. The patches may or may not be topologically contiguous. Within a patch, the actions of agents are coordinated to

maximize the aggregate figure of merit for the entire patch. The size and location of patches are parameters for this strategy.

p A neighborhood is defined for a node such that when a decision is made there, figures of merit at the current node and at a proportion p of neighboring nodes are taken into account. A neighborhood need not consist of the immediate topological neighbors of the node.

tau Only a fraction (called *tau*) of the agents make decisions that change the portions of their state that affect the reward of other agents at the same time.

FIG. 2 provides a flow diagram 200 for determining optimal values of parameters of methods performing routing control and system control. In step 210, the present invention defines a global performance measure for the network. In step 220, the present invention defines an optimization algorithm having at least one parameter. Exemplary parameters include the size and location of patches, the neighborhood, *p* where the figures of merit are considered in making a decision and the fraction, *tau*, of the agents that change portions of their state that affect the reward of other agents. In step 230, the method 200 constructs a landscape representation for values of the parameters and their associated global performance measure. In step 240, the method optimizes over the landscape to produce optimal values for the parameters.

In the preferred embodiment, the present invention uses either patches or *p* or both to define a modified reward and hence, a return, for an agent in the network routing problem. As explained earlier, the figure of merit for an agent is

either its earnings in the market-based framework or its local measure of performance in the local performance framework. Optionally, the present invention uses the *tau* strategy either alone, or in conjunction with *p* and "patches" to limit the opportunities agents have for making decisions that affect the return of other agents. For example, the reward for an agent is the aggregate earnings for a region of agents (a patch) and the bids and routing tables for only a fraction *tau* of agents change at the same time. Preferably, the parameters for these strategies (the fraction p, the fraction *tau* and the number and membership of patches) are global in nature. In other words, the values of these parameters are the same for all agents. Alternatively, the values of the parameters may vary among the agents.

Preferably, the present invention sets these parameters as follows:

First, a global performance measure is defined. Preferably, the global performance measure is a combination of the average delivery time and the achieved network bandwidth. Second, the algorithm has an outer loop that varies these parameters in order to maximize the global performance measure in accordance with techniques for searching landscapes as described in the co-pending international patent application titled, "A System and Method for the Analysis and Prediction of Economic Markets", filed December 22, 1999 at the U.S. receiving office, the contents of which are herein incorporated by reference.

Preferably, each value of the global parameters governing *p*, patches, *tau*, and reinforcement learning features, defines a point in the global parameter space. With respect to this point, the bandwidth-agent system of the present invention achieves a given global fitness. The distribution of global fitness values over the global parameter space constitutes a "fitness landscape" for the

entire bandwidth-agent system. Such landscapes typically have many peaks of high fitness, and statistical features such as correlation lengths and other features as described in co-pending international patent application number PCT/US99/19916, titled, "A Method for Optimal Search on a Technology Landscape", the contents of which are herein incorporated by reference. In the preferred embodiment, these features are used to optimize an evolutionary search in the global parameter space to achieve values of *p*, patches, *tau*, and the internal parameters of the reinforcement learning algorithm that optimize the learning performance of the bandwidth-agent system in a stationary environment with respect to load and other use factor distribution. Preferably, the same search procedures are also used to persistently tune the global parameters of the bandwidth-agent system in a non-stationary environment with respect to load and other use factor distributions.

By tuning of the global parameters to optimize learning, the present invention is "self calibrating". In other words, the invention includes an outer loop in its learning procedure to optimize learning itself, where co-evolutionary learning is in turn controlled by combinations of *p*, patches, and *tau*, plus features of the reinforcement learning algorithm. The inclusion of features of fitness landscapes aids optimal search in this outer loop for global parameter values that themselves optimize learning by the bandwidth-agent system in stationary and non-stationary environments.

Use of *p*, *tau*, or patches aids adaptive search on rugged landscapes because, each by itself, causes the evolving system to ignore some of the constraints some of the time. Judicious balancing of ignoring some of the constraints some of the time with search over the landscape optimizes the balance between "exploitation" and "exploration". In particular, without the capacity to ignore some of the

constraints some of the time, adaptive systems tend to become trapped on local, very sub-optimal peaks. The capacity to ignore some of the constraints some of the time allows the total adapting system to escape badly sub-optimal peaks on the fitness landscape and thereby, enables further searching. In the preferred embodiment, the present invention tunes *p*, tau, or patches either alone or in conjunction with one another to find the proper balance between stubborn exploitation hill climbing and wider exploration search.

The optimal character of either tau alone or patches alone, is such that the total adaptive system is poised slightly in the ordered regime, near a phase transition between order and chaos. *See e.g.* "At Home in the Universe" by Kauffman, Chapters 1,4, 5 and 11, the contents of which are herein incorporated by reference and "The Origins of Order, Stuart Kauffman, Oxford University Press, 1993, Chapters 5 and 6, the contents of which are herein incorporated by reference. For the *p* parameter alone, the optimal value of *p* is not associated with a phase transition.

Without limitation, the embodiments of the present invention are described in the illustrative context of a solution using *tau*, *p*, and patches. However, it will be apparent to persons of ordinary skill in the art that other techniques that ignore some of the constraints some of the time could be used to embody the aspect of the present invention which includes defining an algorithm having one or more parameters, defining a global performance measure, constructing a landscape representation for values of the parameters and their associated global performance value, and optimizing over the landscape to determine optimal values for the parameters. Other exemplary techniques that ignore some of the constraints some of the time include simulated annealing, or optimization at a fixed temperature. In general, the present invention employs the union of any of these means to ignore some of the constraints some of the

time together with reinforcement learning to achieve good problem optimization.

Further, there are local characteristics in the adapting system itself that can be used to test locally that the system is optimizing well. In particular, with patches alone and tau alone, the optimal values of these parameters for adaptation are associated with a power law distribution of small and large avalanches of changes in the system as changes introduced at one point to improve the system unleash a cascade of changes at nearby points in the system. The present invention includes the use of local diagnostics such as a power law distribution of avalanches of change, which are measured either in terms of the size of the avalanches, or in terms of the duration of persistent changes at any single site in the network.

The present invention's use of any combination of the above strategies, together with reinforcement learning in any of its versions, give it an advantage over prior art routing methods because these strategies address many problems that could arise including the following:

- slow convergence to optimal routing patterns,
- oscillation of network load, and
- locally beneficial but globally harmful routing patterns.

Without limitation, the embodiments of the present invention have been described in the illustrative context of a method for routing data through a communication network. However, it is apparent to persons of ordinary skill in the art that other contexts could be used to embody the aspect of the present invention which includes defining an algorithm having one or more parameters, defining a global performance measure, constructing a landscape representation for values of the parameters and their associated global performance

value, and optimizing over the landscape to determine optimal values for the parameters.

For example, the present invention could be used for operations management as explained in co-pending U.S. patent application No. 09/345,441, titled, "An Adaptive and Reliable System and Method for Operations management" and filed on July 1, 1999, the contents of which are herein incorporated by reference. That patent describes a model of an enterprise in its competitive environment, based on technology graphs that support a nodes and flow model of an organization, plus a management structure. The present invention, using agents to represent objects and operations in the enterprise model, coupled to reinforcement learning, *p*, patches and *tau*, is used advantageously to create a model of a learning organization that learns how to adapt well in its local environment. By use of the outer loop described above, good global parameter values for *p*, patches, *tau*, and the reinforcement learning algorithm are discovered. In turn, these values are used to help create homologous action patterns in the real organization. For example, the homologous action patters can be created by tuning the partitioning the organization into patches, by tuning how decisions at one point in the real organization are taken with respect to a prospective benefit of a fraction *p* of the other points in the organization affected by the first point, and by tuning what fraction, *tau*, of points in the organization should try operational and other experiments to improve performance.

In addition, the distribution of contract values and rewards in the reinforcement algorithm can be used to help find good incentive structures to mediate behavior by human agents in the real organization to achieve the overall adaptive and agile performance of the real organization.

In addition to the use of the invention to find good global parameters to instantiate in the real organization, the same invention can be used to find good global parameter values to utilize in the model of the organization itself to use that model as a decision support tool, teaching tool, etc.

Further, the present invention is also applicable to portfolio management, risk management, scheduling and routing problems, logistic problems, supply chain problems and other practical problems characterized by many interacting factors.

While the above invention has been described with reference to certain preferred embodiments, the scope of the present invention is not limited to these embodiments. One skill in the art may find variations of these preferred embodiments which, nevertheless, fall within the spirit of the present invention, whose scope is defined by the claims set forth below.

Claims

Claims

1. A method for routing packets of data through a network of a plurality of components comprising the steps of:

controlling one or more of said components by executing a corresponding one or more software agents, comprising the steps of:

receiving information for at least one of the packets;

computing an expected return for delivery of said at least one packet from said information; and

directing the delivery of said at least one packet to optimize said expected return.

2. A method as in claim 1 wherein said information for said at least one packet comprises a destination.

3. A method as in claim 2 wherein said information for said at least one packet further comprises a contract to pay a specified reward to said one or more software agents that delivers said at least one packet to said destination.

4. A method as in claim 3 wherein said information of said at least one packet further comprises a specified quality of service.

5. A method as in claim 4 wherein said specified reward varies with a delivered quality of service in comparison with said specified quality of service.

6. A method as in claim 4 wherein said information for said at least one packet comprises at least one bid specifying a price that said one or more software agent will pay for said at least one packet having said destination and said quality of service.

7. A method as in claim 4 wherein said quality of service comprises a latency for said at least one packet.

8. A method as in claim 4 wherein said quality of service comprises a specified order for delivery of said at least one packet.

9. A method as in claim 1 wherein said information for said at least one packet comprises at least one bid specifying a price that said one or more software agent will pay for said at least one packet.

10. A method as in claim 9 wherein said at least one bid further comprises an expiration time.

11. A method as in claim 9 wherein said at least one bid further comprises a margin.

12. A method as in claim 9 wherein said at least one bid further comprises a minimum value.

13. A method as in claim 9 wherein said at least one bid further comprises a minimum delivery time.

14. A method as in claim 9 wherein said at least one bid further comprises a specified route.

15. A method as in claim 9 wherein said at least one bid is a satisfaction profile defining a satisfaction of trading said at least one packet as a probability density function of at least one parameter.

16. A method as in claim 15 wherein said at least one parameter of said probability density function comprises a quality of service.

17. A method as in claim 1 wherein said expected return for delivery of said at least one packet is an expected reward discounted to present value.

18. A method as in claim 1 wherein said expected return for delivery of said at least one packet step varies inversely with an estimated delivery time for said at least one packet.

19. A method as in claim 18 wherein said controlling one or more components step further comprises the step of transmitting delivery loads to others of said one or more software agents for determining said estimated delivery time for said at least one packet.

20. A method as in claim 1 wherein said one or more software agents control one or more legal entities of the network.

21. A method as in claim 1 wherein said one or more software agents control one or more communication links of the network.

22. A method as in claim 1 wherein said controlling one or more of said components step further

comprises the step of partitioning said one or more software agents into one or more patches.

23. A method as in claim 22 wherein said directing the delivery of said at least one packet step comprises the step of optimizing said expected return of said patch.

24. A method as in claim 1 wherein said computing an expected return step comprises the step of:

selecting a portion p of said one or more software agents; and

computing said expected return of said selected portion p of said one or more software agents.

25. A method as in claim 24 wherein said delivery of said at least one packet is directed to optimize said expected return of said selected portion p of said one or more software agents.

26. A method as in claim 1 wherein said controlling one or more of said components step further comprises the step of transmitting said information from said one or more software agents to others of said software agents.

27. A method as in claim 26 wherein said transmitted information comprises at least one bid specifying a price that said one or more software agents will pay for said at least one packet.

28. A method as in claim 26 wherein said transmitted information comprise delivery loads.

29. A method as in claim 26 wherein only a fraction, tau, of said one or more software agents transmit said information at the same time.

30. A method for routing packets of data through a network of components comprising the steps of:

defining at least one algorithm having one or more parameters for routing the data;

defining at least one global performance measure of said at least one algorithm;

executing said algorithm for a plurality of different values of said one or more parameters to generate a corresponding plurality of values for said global performance measure;

constructing a fitness landscape from said values of said parameters and said corresponding values of said global performance measure; and

optimizing over said fitness landscape to generate optimal values for said at least one parameter.

31. A method as in claim 30 wherein said defining an algorithm step comprises the steps of:

controlling one or more of said components by executing a corresponding one or more software agents comprising the steps of:

communicating information for at least one of the packets among said one or more software agents;

computing an expected return for delivery of said at least one packet from said information; and

directing the delivery of said at least one packet to optimize said expected return.

32. A method as in claim 31 wherein said at least one parameter comprises a proportion p of said one or more software agents.

33. A method as in claim 32 wherein said computing an expected return step comprises the step of:

computing said expected return of said proportion p of said one or more software agents.

34. A method as in claim 31 wherein said at least one parameter comprises a size of one or more patches of said one or more software agents and a location of said patches.

35. A method as in claim 34 wherein said directing the delivery of said at least one packet step comprises the step of:

optimizing said expected return of said patch.

36. A method as in claim 31 wherein said at least one parameter comprises a fraction, tau, of said one or more software agents.

37. A method as in claim 36 wherein only said fraction, tau, of said software agents communicate information for said at least one packet at the same time.

38. A method for performing operations management in an environment of entities comprising the steps of:

representing at least one of the entities with at least one corresponding model having a plurality of parameters;

defining at least one global performance measure of said model;

executing said model for a plurality of different values of said at least one parameters to generate a corresponding plurality of values for said global performance measure;

constructing a fitness landscape from said values of said parameters and said corresponding values of said global performance measure; and

optimizing over said fitness landscape to generate optimal values for said at least one parameter.

39. A method as in claim 38 wherein said representing at least one of the entities with at least one corresponding model having a plurality of parameters step comprises the steps of:

representing a plurality of decision making units within the entities with a corresponding plurality of decision making agents; and

representing a plurality of communication links among the decision making units with a corresponding plurality of connections among said plurality of decision making agents.

40. A method as in claim 39 further comprising the steps of:

communicating information among said decision making agents;

computing an expected return at said decision making agents from said information; and

making at least one decision at said decision making agent to optimize said expected return.

41. A method as in claim 40 wherein said at least one parameter comprises a proportion p of said decision making agents.

42. A method as in claim 41 wherein said computing an expected return step comprises the step of:

computing said expected return of said proportion *p* of said decision making agents.

43. A method as in claim 40 wherein said at least one parameter comprises a size of one or more patches of said decision making agents and a location of said patches.

44. A method as in claim 43 wherein said making at least one decision step comprises the step of:

optimizing said expected return of said patch.

45. A method as in claim 40 wherein said at least one parameter comprises a fraction, *tau*, of said decision making agents.

46. A method as in claim 45 wherein only said fraction, tau, of said decision making agents communicate information at the same time.

47. Computer executable software code stored on a computer readable medium, the code for routing packets of data through a network of a plurality of components, the code comprising:

code to control one or more of said components by executing a corresponding one or more software agents, comprising:

code to receive information for at least one of the packets;

code to compute an expected return for delivery of said at least one packet from said information; and

code to direct the delivery of said at least one packet to optimize said expected return.

50. A programmed component for routing packets of data through a network comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory, wherein the program code comprises:

code to control one or more of said components by executing a corresponding one or more software agents, comprising:

code to receive information for at least one of the packets;

code to compute an expected return for delivery of said at least one packet from said information; and

code to direct the delivery of said at least one packet to optimize said expected return.

49. Computer executable software code stored on a computer readable medium, the code for routing packets of data through a network of a plurality of components, the code comprising:

code to define at least one algorithm having one or more parameters for routing the data;

code to define at least one global performance measure of said at least one algorithm;

code to execute said algorithm for a plurality of different values of said one or more parameters to generate a corresponding plurality of values for said global performance measure;

code to construct a fitness landscape from said values of said parameters and said corresponding values of said global performance measure; and

code to optimize over said fitness landscape to generate optimal values for said at least one parameter.

50. A programmed component for routing packets of data through a network comprising at least one memory having at least one region storing computer executable program code and at least one processor for executing the program code stored in said memory, wherein the program code comprises:

code to define at least one algorithm having one or more parameters for routing the data;

code to define at least one global performance measure of said at least one algorithm;

code to execute said algorithm for a plurality of different values of said one or more parameters to generate a corresponding plurality of values for said global performance measure;

code to construct a fitness landscape from said values of said parameters and said corresponding values of said global performance measure; and

code to optimize over said fitness landscape to generate optimal values for said at least one parameter.

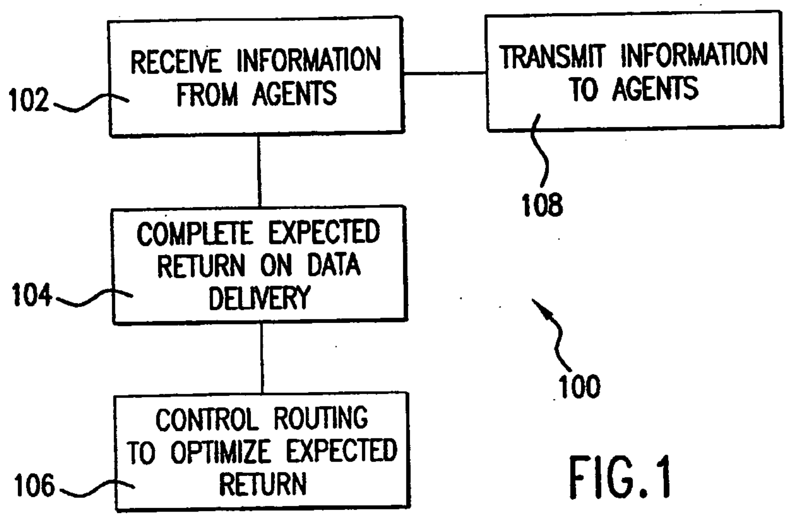

RECEIVE INFORMATION FROM AGENTS
102
TRANSMIT INFORMATION TO AGENTS
108
COMPLETE EXPECTED RETURN ON DATA DELIVERY
104
100
CONTROL ROUTING TO OPTIMIZE EXPECTED RETURN
106

FIG.1

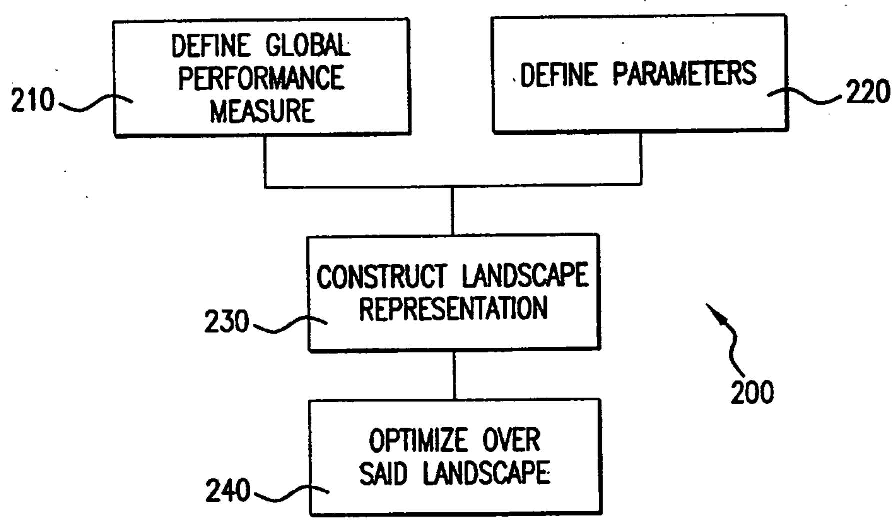

DEFINE GLOBAL PERFORMANCE MEASURE
210
DEFINE PARAMETERS
220
CONSTRUCT LANDSCAPE REPRESENTATION
230
200
OPTIMIZE OVER SAID LANDSCAPE
240

FIG.2

INTERNATIONAL SEARCH REPORT	International application No. PCT/US00/02011

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04M 15/00; H04L 12/46; G06F 15/18
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/252, 253, 255, 270, 400, 408; 706/8, 12, 19, 25; 709/238, 242, 243

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,E --- Y	US 6,046,985 A (ALDRED ET AL) 04 APRIL 2000, col. 4, line 59 to col. 11, line 39.	1, 2, 20, 21, 26, 28 ------- 3-16, 27, 47-48
Y	US 5,790,642 A (TAYLOR ET AL) 04 AUGUST 1998, col. 4, line 11 to col. 9, line 53.	3-16, 27, 47-48
A	AHN ET AL, GENROUTER: A GENETIC ALGORITHM FOR CHANNEL ROUTING PROBLEMS, IEEE, MAY 1995, pp. 151-154.	1-50

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 JUNE 2000	06 JUL 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DUONG, FRANK Joni Hill Telephone No. (703) 308-5428

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/02011

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

370/252, 253, 255, 270, 400, 408; 706/8, 12, 19, 25; 709/238, 242, 243